201 AGACCGGGCGGCGGCTTTGGATTTTGGGGGGGCGGGGACCAGCTGCGCGG 250

251 CGGCACCATGTTCCTAGCCACTCTGTACTTCGCGCTGCCACTCCTGGATT 300
   M F L A T L Y F A L P L L D L

301 TGCTGATGTCCGCCGAGGTGAGTGGTGGAGACCGTCTGGACTGTGTGAAA 350
  L M S A E V S G G D R L D C V K

351 GCCAGCGATCAGTGCCTGAAGGAACAGAGCTGCAGCACCAAGTACCGCAC 400
  A S D Q C L K E Q S C S T K Y R T

401 ACTAAGGCAGTGCGTGGCGGGCAAGGAAACCAACTTCAGCCTGACATCCG 450
  L R Q C V A G K E T N F S L T S G

451 GCCTTGAGGCCAAGGATGAGTGCCGTAGCGCCATGGAGGCCTTGAAGCAG 500
  L E A K D E C R S A M E A L K Q

501 AAGTCTCTGTACAACTGCCGCTGCAAGCGGGGCATGAAGAAAGAGAAGAA 550
  K S L Y N C R C K R G M K K E K N

551 TTGTCTGCGTATCTACTGGAGCATGTACCAGAGCCTGCAGGGAAATGACC 600
  C L R I Y W S M Y Q S L Q G N D L

601 TCCTGGAAGATTCCCCGTATGAGCCGGTTAACAGCAGGTTGTCAGATATA 650
  L E D S P Y E P V N S R L S D I

651 TTCCGGGCAGTCCCGTTCATATCAGATGTTTTCCAGCAAGTGGAACACAT 700
  F R A V P F I S D V F Q Q V E H I

701 TTCCAAAGGGAACAACTGCCTGGACGCAGCCAAGGCCTGCAACCTGGACG 750
  S K G N N C L D A A K A C N L D D

751 ACACCTGTAAGAAGTACAGGTCGGCCTACATCACCCCCTGCACCACCAGC 800
  T C K K Y R S A Y I T P C T T S

801 ATGTCCAACGAGGTCTGCAACCGCCGTAAGTGCCACAAGGCCCTCAGGCA 850
  M S N E V C N R R K C H K A L R Q

851 GTTCTTCGACAAGGTTCCGGCCAAGCACAGCTACGGGATGCTCTTCTGCT 900
  F F D K V P A K H S Y G M L F C S

901 CCTGCCGGGACATCGCCTGCACCGAGCGGCGGCGACAGACTATCGTCCCC 950
  C R D I A C T E R R R Q T I V P

951 GTGTGCTCCTATGAAGAACGAGAGAGGCCCAACTGCCTGAGTCTGCAAGA 1000
  V C S Y E E R E R P N C L S L Q D

1001 CTCCTGCAAGACCAATTACATCTGCAGATCTCGCCTTGCAGATTTTTTTA 1050
  S C K T N Y I C R S R L A D F F T

1051 CCAACTGCCAGCCAGAGTCAAGGTCTGTCAGCAACTGTCTTAAGGAGAAC 1100
  N C Q P E S R S V S N C L K E N

1101 TACGCAGACTGCCTCCTGGCCTACTCGGGACTGATTGGCACAGTCATGAC 1150
  Y A D C L L A Y S G L I G T V M T

FIG. 1a

```
1151  TCCCAACTACGTAGACTCCAGCAGCCTCAGCGTGGCACCATGGTGTGACT  1200
        P  N  Y  V  D  S  S  S  L  S  V  A  P  W  C  D  C

1201  GCAGCAACAGCGGCAATGACCTGGAAGACTGCTTGAAATTTCTGAATTTT  1250
        S  N  S  G  N  D  L  E  D  C  L  K  F  L  N  F

1251  TTTAAGGACAATACTTGTCTCAAAAATGCAATTCAAGCCTTTGGCAATGG  1300
        F  K  D  N  T  C  L  K  N  A  I  Q  A  F  G  N  G

1301  CTCAGATGTGACCATGTGGCAGCCAGCCCCTCCAGTCCAGACCACCACTG  1350
        S  D  V  T  M  W  Q  P  A  P  P  V  Q  T  T  T  A

1351  CCACCACTACCACTGCCTTCCGGGTCAAGAACAAGCCTCTGGGGCCAGCA  1400
        T  T  T  T  A  F  R  V  K  N  K  P  L  G  P  A

1401  GGGTCTGAGAATGAGATCCCCACACACGTTTTACCACCCTGTGCGAATTT  1450
        G  S  E  N  E  I  P  T  H  V  L  P  P  C  A  N  L

1451  GCAGGCTCAGAAGCTGAAATCCAATGTGTCGGGTAGCACACACCTCTGTC  1500
        Q  A  Q  K  L  K  S  N  V  S  G  S  T  H  L  C  L

1501  TTTCTGATAGTGATTTCGGAAAGGATGGTCTCGCTGGTGCCTCCAGCCAC  1550
        S  D  S  D  F  G  K  D  G  L  A  G  A  S  S  H

1551  ATAACCACAAAATCAATGGCTGCTCCTCCCAGCTGCAGTCTGAGCTCACT  1600
        I  T  T  K  S  M  A  A  P  P  S  C  S  L  S  S  L

1601  GCCGGTGCTGATGCTCACCGCCCTTGCTGCCCTGTTATCTGTATCGTTGG  1650
        P  V  L  M  L  T  A  L  A  A  L  L  S  V  S  L  A

1651  CAGAAACGTCGTAGCTGCATCCGGGAAAACAGTATGAAAAGACAAAAGAG  1700
        E  T  S
```

FIG. 1b

```
  1   CTGCTGGAGGATTCCCCATATGAACCAGTTAACAGCAGATTGTCAGATAT   50
      L  L  E  D  S  P  Y  E  P  V  N  S  R  L  S  D  I

 51   ATTCCGGGTGGTCCCATTCATATCAGTGGAGCACATTCCCAAAGGGAACA   100
      F  R  V  V  P  F  I  S  V  E  H  I  P  K  G  N  N

101   ACTGCCTGGATGCAGCGAAGGCCTGCAACCTCGACGACATTTGCAAGAAG   150
      C  L  D  A  A  K  A  C  N  L  D  D  I  C  K  K

151   TACAGGTCGGCGTACATCACCCCGTGCACCACCAGCGTGTCCAACGATGT   200
      Y  R  S  A  Y  I  T  P  C  T  T  S  V  S  N  D  V

201   CTGCAACCGCCGCAAGTGCCACAAGGCCCTCCGGCAGTTCTTTGACAAGG   250
      C  N  R  R  K  C  H  K  A  L  R  Q  F  F  D  K  V

251   TCCCGGCCAAGCACAGCTACGGAATGCTCTTCTGCTCCTGCCGGGACATC   300
      P  A  K  H  S  Y  G  M  L  F  C  S  C  R  D  I

301   GCCTGCACAGAGCGGAGGCGACAGACCATCGTGCCTGTGTGCTCCTATGA   350
      A  C  T  E  R  R  R  Q  T  I  V  P  V  C  S  Y  E

351   AGAGAGGGAGAAGCCCAACTGTTTGAATTTGCAGGACTCCTGCAAGACGA   400
      E  R  E  K  P  N  C  L  N  L  Q  D  S  C  K  T  N

401   ATTACATCTGCAGATCTCGCCTTGCGGATTTTTTTACCAACTGCCAGCCA   450
      Y  I  C  R  S  R  L  A  D  F  F  T  N  C  Q  P

451   GAGTCAAGGTCTGTCAGCAGCTGTCTAAAGGAAAACTACGCTGACTGCCT   500
      E  S  R  S  V  S  S  C  L  K  E  N  Y  A  D  C  L

501   CCTCGCCTACTCGGGGCTTATTGGCACAGTCATGACCCCCAACTACATAG   550
      L  A  Y  S  G  L  I  G  T  V  M  T  P  N  Y  I  D

551   ACTCCAGTAGCCTCAGTGTGGCCCCATGGTGTGACTGCAGCAACAGTGGG   600
      S  S  S  L  S  V  A  P  W  C  D  C  S  N  S  G

601   AACGACCTAGAAGAGTGCTTGAAATTTTTGAATTTCTTCAAGGACAATAC   650
      N  D  L  E  E  C  L  K  F  L  N  F  F  K  D  N  T

651   ATGTCTTAAAAATGCAATTCAAGCCTTTGGCAATGGCTCCGATGTGACCG   700
      C  L  K  N  A  I  Q  A  F  G  N  G  S  D  V  T  V

701   TGTGGCAGCCAGCCTTCCCAGTACAGACCACCACTGCCACTACCACCACT   750
      W  Q  P  A  F  P  V  Q  T  T  T  A  T  T  T  T

751   GCCCTCCGGGTTAAGAACAAGCCCCTGGGGCCAGCAGGGTCTGAGAATGA   800
      A  L  R  V  K  N  K  P  L  G  P  A  G  S  E  N  E

801   AATTCCCACTCATGTTTTGCCACCGTGTGCAAATTTACAGGCACAGAAGC   850
      I  P  T  H  V  L  P  P  C  A  N  L  Q  A  Q  K  L

851   TGAAATCCAATGTGTCGGGCAATACACACCTCTGTATTTCCAATGGTAAT   900
      K  S  N  V  S  G  N  T  H  L  C  I  S  N  G  N
```

FIG. 2A-1

```
901   TATGAAAAAGAAGGTCTCGGTGCTTCCAGCCACATAACCACAAAATCAAT   950
      Y  E  K  E  G  L  G  A  S  S  H  I  T  T  K  S  M

951   GGCTGCTCCTCCAAGCTGTGGTCTGAGCCCACTGCTGGTCCTGGTGGTAA   1000
      A  A  P  P  S  C  G  L  S  P  L  L  V  L  V  V  T

1001  CCGCTCTGTCCACCCTATTATCTTTAACAGAACATCATAGCTGCATTAA   1050
      A  L  S  T  L  L  S  L  T  E  T  S

1051  AAAAATACAATATGGACATGTAAAAGACAAAAACCAAGTTATCTGTTTC   1100

1101  CTGTTCTCTTGTATAGCTGAAATTCCAGTTTAGGAGCTCAGTTGAGAAAC   1150

1151  AGTTCCATTCAACTGGAACATTTTTTTTTTTTCCTTTTAAGAAAGCTTCT   1200

1201  TGTGATCCTTCGGGGCTTCTGTG   1223
```

FIG. 2A-2

```
  1  GGGCGGCCAGAGCAGCACAGCTGTCCGGGGATCGCTGCATGCTGAGCTCC      50

 51  CTCGGCAAGACCCAGCGGCGGCTCGGGATTTTTTTGGGGGGGCGGGGACC     100

101  AGCCCCGCGCCGGCACCATGTTCCTGGCGACCCTGTACTTCGCGCTGCCG     150
                        M  F  L  A  T  L  Y  F  A  L  P

151  CTCTTGGACTTGCTCCTGTCGGCCGAAGTGAGCGGCGGAGACCGCCTGGA     200
      L  L  D  L  L  L  S  A  E  V  S  G  G  D  R  L  D

201  TTGCGTGAAAGCCAGTGATCAGTGCCTGAAGGAGCAGAGCTGCAGCACCA     250
      C  V  K  A  S  D  Q  C  L  K  E  Q  S  C  S  T  K

251  AGTACCGCACGCTAAGGCAGTGCGTGGCGGGCAAGGAGACCAACTTCAGC     300
      Y  R  T  L  R  Q  C  V  A  G  K  E  T  N  F  S

301  CTGGCATCCGGCCTGGAGGCCAAGGATGAGTGCCGCAGCGCCATGGAGGC     350
      L  A  S  G  L  E  A  K  D  E  C  R  S  A  M  E  A

351  CCTGAAGCAGAAGTCGCTCTACAACTGCCGCTGCAAGCGGGGTATGAAGA     400
      L  K  Q  K  S  L  Y  N  C  R  C  K  R  G  M  K  K

401  AGGAGAAGAACTGCCTGCGCATTTACTGGAGCATGTACCAGAGCCTGCAG     450
      E  K  N  C  L  R  I  Y  W  S  M  Y  Q  S  L  Q

451  GGAAATGATCTGCTGGAGGATTCCCCATATGAACCAGTTAACAGCAGATT     500
      G  N  D  L  L  E  D  S  P  Y  E  P  V  N  S  R  L

501  GTCAGATATATTCCGGGTGGTCCCATTCATATCAGTGGAGCACATTCCCA     550
      S  D  I  F  R  V  V  P  F  I  S  V  E  H  I  P  K

551  AAGGGAACAACTGCCTGGATGCAGCGAAGGCCTGCAACCTCGACGACATT     600
      G  N  N  C  L  D  A  A  K  A  C  N  L  D  D  I

601  TGCAAGAAGTACAGGTCGGCGTACATCACCCCGTGCACCACCAGCGTGTC     650
      C  K  K  Y  R  S  A  Y  I  T  P  C  T  T  S  V  S

651  CAACGATGTCTGCAACCGCCGCAAGTGCCACAAGGCCCTCCGGCAGTTCT     700
      N  D  V  C  N  R  R  K  C  H  K  A  L  R  Q  F  F

701  TTGACAAGGTCCCGGCCAAGCACAGCTACGGAATGCTCTTCTGCTCCTGC     750
      D  K  V  P  A  K  H  S  Y  G  M  L  F  C  S  C

751  CGGGACATCGCCTGCACAGAGCGGAGGCGACAGACCATCGTGCCTGTGTG     800
      R  D  I  A  C  T  E  R  R  R  Q  T  I  V  P  V  C

801  CTCCTATGAAGAGAGGGAGAAGCCCAACTGTTTGAATTTGCAGGACTCCT     850
      S  Y  E  E  R  E  K  P  N  C  L  N  L  Q  D  S  C

851  GCAAGACGAATTACATCTGCAGATCTCGCCTTGCGGATTTTTTTACCAAC     900
      K  T  N  Y  I  C  R  S  R  L  A  D  F  F  T  N

901  TGCCAGCCAGAGTCAAGGTCTGTCAGCAGCTGTCTAAAGGAAAACTACGC     950
      C  Q  P  E  S  R  S  V  S  S  C  L  K  E  N  Y  A
```

FIG. 2B-1

951 TGACTGCCTCCTCGCCTACTCGGGGCTTATTGGCACAGTCATGACCCCCA 1000
D   C   L   L   A   Y   S   G   L   I   G   T   V   M   T   P   N

1001 ACTACATAGACTCCAGTAGCCTCAGTGTGGCCCCATGGTGTGACTGCAGC 1050
Y   I   D   S   S   S   L   S   V   A   P   W   C   D   C   S

1051 AACAGTGGGAACGACCTAGAAGAGTGCTTGAAATTTTTGAATTTCTTCAA 1100
N   S   G   N   D   L   E   E   C   L   K   F   L   N   F   F   K

1101 GGACAATACATGTCTTAAAAATGCAATTCAAGCCTTTGGCAATGGCTCCG 1150
D   N   T   C   L   K   N   A   I   Q   A   F   G   N   G   S   D

1151 ATGTGACCGTGTGGCAGCCAGCCTTCCCAGTACAGACCACCACTGCCACT 1200
V   T   V   W   Q   P   A   F   P   V   Q   T   T   T   A   T

1201 ACCACCACTGCCCTCCGGGTTAAGAACAAGCCCCTGGGGCCAGCAGGGTC 1250
T   T   T   A   L   R   V   K   N   K   P   L   G   P   A   G   S

1251 TGAGAATGAAATTCCCACTCATGTTTTGCCACCGTGTGCAAATTTACAGG 1300
E   N   E   I   P   T   H   V   L   P   P   C   A   N   L   Q   A

1301 CACAGAAGCTGAAATCCAATGTGTCGGGCAATACACACCTCTGTATTTCC 1350
Q   K   L   K   S   N   V   S   G   N   T   H   L   C   I   S

1351 AATGGTAATTATGAAAAGAAGGTCTCGGTGCTTCCAGCCACATAACCAC 1400
N   G   N   Y   E   K   E   G   L   G   A   S   S   H   I   T   T

1401 AAAATCAATGGCTGCTCCTCCAAGCTGTGGTCTGAGCCCACTGCTGGTCC 1450
K   S   M   A   A   P   P   S   C   G   L   S   P   L   L   V   L

1451 TGGTGGTAACCGCTCTGTCCACCCTATTATCTTTAACAGAACATCATAG 1500
V   V   T   A   L   S   T   L   L   S   L   T   E   T   S

1501 CTGCATTAAAAAAATACAATATGGACATGTAAAAAGACAAAAACCAAGTT 1550

1551 ATCTGTTTCCTGTTCTCTTGTATAGCTGAAATTCCAGTTTAGGAGCTCAG 1600

1601 TTGAGAAACAGTTCCATTCAACTGGAACATTTTTTTTTTTTCCTTTTAAG 1650

1651 AAAGCTTCTTGTGATCCTTCGGGGCTTCTGTG 1682

FIG. 2B-2

```
  1 GGGCGGCCAGAGCAGCACAGCTGTCCGGGGATCGCTGCATGCTGAGCTCC  50
    | | |||||||||| ||| |||  |  ||||||||||||||||  ||||||||||
143 GAGTGGCCAGAGGAGCGCAGTCGCCCGGGGATCGCTGCACGCTGAGCTCT 192

 51 CTCGGCAAGACCCAGCGGCGGCTCGGGATTTTTTTGGGGGGGCGGGGACC 100
    ||| | |||||| ||||||||||| |||  |||||||||||||||||||||
193 CTCCCCGAGACCGGGCGGCGGCTTTGGA...TTTTGGGGGGGCGGGGACC 239

101 AGCCCCGCGCCGGCACCATGTTCCTGGCGACCCTGTACTTCGCGCTGCCG 150
    ||| ||||| |||||||||||||||||||| || || |||||||||||||||
240 AGCTGCGCGGCGGCACCATGTTCCTAGCCACTCTGTACTTCGCGCTGCCA 289

151 CTCTTGGACTTGCTCCTGTCGGCCGAAGTGAGCGGCGGAGACCGCCTGGA 200
    ||| |||| ||||| |||| |||| |||| ||||| |||||||||| |||||
290 CTCCTGGATTTGCTGATGTCCGCCGAGGTGAGTGGTGGAGACCGTCTGGA 339

201 TTGCGTGAAAGCCAGTGATCAGTGCCTGAAGGAGCAGAGCTGCAGCACCA 250
    || |||||||||||| |||||||||||||||||||| ||||||||||||||
340 CTGTGTGAAAGCCAGCGATCAGTGCCTGAAGGAACAGAGCTGCAGCACCA 389

251 AGTACCGCACGCTAAGGCAGTGCGTGGCGGGCAAGGAGACCAACTTCAGC 300
    |||||||| |||||||||||||||||||||||||||||||| ||||||||||
390 AGTACCGCACACTAAGGCAGTGCGTGGCGGGCAAGGAAACCAACTTCAGC 439

301 CTGGCATCCGGCCTGGAGGCCAAGGATGAGTGCCGCAGCGCCATGGAGGC 350
    ||| |||||||||| ||||||||||||||||||||||| ||||||||||||
440 CTGACATCCGGCCTTGAGGCCAAGGATGAGTGCCGTAGCGCCATGGAGGC 489

351 CCTGAAGCAGAAGTCGCTCTACAACTGCCGCTGCAAGCGGGGTATGAAGA 400
    | ||||||||||||| || ||||||||||||||||||||||||||| |||||
490 CTTGAAGCAGAAGTCTCTGTACAACTGCCGCTGCAAGCGGGGCATGAAGA 539

401 AGGAGAAGAACTGCCTGCGCATTTACTGGAGCATGTACCAGAGCCTGCAG 450
    | ||||||||| || |||||| || |||||||||||||||||||||||||||
540 AAGAGAAGAATTGTCTGCGTATCTACTGGAGCATGTACCAGAGCCTGCAG 589

451 GGAAATGATCTGCTGGAGGATTCCCCATATGAACCAGTTAACAGCAGATT 500
    ||||||||| || |||||| |||||||||||| || |||||||||||| ||
590 GGAAATGACCTCCTGGAAGATTCCCCGTATGAGCCGGTTAACAGCAGGTT 639

501 GTCAGATATATTCCGGGTGGTCCCATTCATATC...............AG 535
    |||||||||||||||||| |||||  |||||||||             ||
640 GTCAGATATATTCCGGGCAGTCCCGTTCATATCAGATGTTTTCCAGCAAG 689

536 TGGAGCACATTCCCAAAGGGAACAACTGCCTGGATGCAGCGAAGGCCTGC 585
    |||| |||||||| ||||||||||||||||||||| ||||| |||||||||
690 TGGAACACATTTCCAAAGGGAACAACTGCCTGGACGCAGCCAAGGCCTGC 739

586 AACCTCGACGACATTTGCAAGAAGTACAGGTCGGCGTACATCACCCCGTG 635
    ||||| |||||||| || |||||||||||||||||| |||||||||||| ||
740 AACCTGGACGACACCTGTAAGAAGTACAGGTCGGCCTACATCACCCCCTG 789

636 CACCACCAGCGTGTCCAACGATGTCTGCAACCGCCGCAAGTGCCACAAGG 685
    ||||||||||| |||||||||| |||||||||||||||||| |||||||||
790 CACCACCAGCATGTCCAACGAGGTCTGCAACCGCCGTAAGTGCCACAAGG 839
```

FIG. 3A-1

```
 686  CCCTCCGGCAGTTCTTTGACAAGGTCCCGGCCAAGCACAGCTACGGAATG  735
      ||||| |||||||||| |||||||| |||||||||||||||||||||| |||
 840  CCCTCAGGCAGTTCTTCGACAAGGTTCCGGCCAAGCACAGCTACGGGATG  889

 736  CTCTTCTGCTCCTGCCGGGACATCGCCTGCACAGAGCGGAGGCGACAGAC  785
      ||||||||||||||||||||||||||||||||||||||| |||||| |||||||||||
 890  CTCTTCTGCTCCTGCCGGGACATCGCCTGCACCGAGCGGCGGCGACAGAC  939

 786  CATCGTGCCTGTGTGCTCCTATGAAGAGAGGGAGAAGCCCAACTGTTTGA  835
      ||||| || |||||||||||||||||| | |||| ||||||||||  |||
 940  TATCGTCCCCGTGTGCTCCTATGAAGAACGAGAGAGGCCCAACTGCCTGA  989

 836  ATTTGCAGGACTCCTGCAAGACGAATTACATCTGCAGATCTCGCCTTGCG  885
      | |||| ||||||||||||||||| |||||||||||||||||||||||||
 990  GTCTGCAAGACTCCTGCAAGACCAATTACATCTGCAGATCTCGCCTTGCA  1039

 886  GATTTTTTACCAACTGCCAGCCAGAGTCAAGGTCTGTCAGCAGCTGTCT  935
      ||||||||||||||||||||||||||||||||||||||||||||| ||||||
1040  GATTTTTTACCAACTGCCAGCCAGAGTCAAGGTCTGTCAGCAACTGTCT  1089

 936  AAAGGAAAACTACGCTGACTGCCTCCTCGCCTACTCGGGGCTTATTGGCA  985
      ||||| ||||||| ||||||||||||||| |||||||||| || ||||||
1090  TAAGGAGAACTACGCAGACTGCCTCCTGGCCTACTCGGGACTGATTGGCA  1139

 986  CAGTCATGACCCCCAACTACATAGACTCCAGTAGCCTCAGTGTGGCCCCA  1035
      |||||||||| ||||||||| |||||||||| |||||| |||| |||
1140  CAGTCATGACTCCCAACTACGTAGACTCCAGCAGCCTCAGCGTGGCACCA  1189

1036  TGGTGTGACTGCAGCAACAGTGGGAACGACCTAGAAGAGTGCTTGAAATT  1085
      ||||||||||||||||||| || || ||||| ||||| |||||||||||
1190  TGGTGTGACTGCAGCAACAGCGGCAATGACCTGGAAGACTGCTTGAAATT  1239

1086  TTTGAATTTCTTCAAGGACAATACATGTCTTAAAAATGCAATTCAAGCCT  1135
      | |||||| || |||||||||||| ||||| |||||||||||||||||||
1240  TCTGAATTTTTTTAAGGACAATACTTGTCTCAAAAATGCAATTCAAGCCT  1289

1136  TTGGCAATGGCTCCGATGTGACCGTGTGGCAGCCAGCCTTCCCAGTACAG  1185
      |||||||||||||| ||||||||| ||||||||||||||   ||||| |||
1290  TTGGCAATGGCTCAGATGTGACCATGTGGCAGCCAGCCCCTCCAGTCCAG  1339

1186  ACCACCACTGCCACTACCACCACTGCCCTCCGGGTTAAGAACAAGCCCCT  1235
      ||||||||||||||||| || |||||||| |||||||| |||||||||| ||
1340  ACCACCACTGCCACCACTACCACTGCCTTCCGGGTCAAGAACAAGCCTCT  1389

1236  GGGGCCAGCAGGGTCTGAGAATGAAATTCCCACTCATGTTTTGCCACCGT  1285
      ||||||||||||||||||||||||| || ||||| || ||||| ||||| |
1390  GGGGCCAGCAGGGTCTGAGAATGAGATCCCCACACACGTTTTACCACCCT  1439

1286  GTGCAAATTTACAGGCACAGAAGCTGAAATCCAATGTGTCGGGCAATACA  1335
      |||| ||||| |||||| |||||||||||||||||||||||||| | |||
1440  GTGCGAATTTGCAGGCTCAGAAGCTGAAATCCAATGTGTCGGGTAGCACA  1489

1336  CACCTCTGTATTTCCAATGGTAATTATGAAAAGAAGGTCTC...GGTGC  1382
      |||||||||| ||||| || || || |||| | |||| || |||||   |||||
1490  CACCTCTGTCTTTCTGATAGTGATTTCGGAAAGGATGGTCTCGCTGGTGC  1539
```

FIG. 3A-2

```
1383 TTCCAGCCACATAACCACAAAATCAATGGCTGCTCCTCCAAGCTGTGGTC 1432
     |||||||||||||||||||||||||||||||||||||||||| ||||| |||
1540 CTCCAGCCACATAACCACAAAATCAATGGCTGCTCCTCCCAGCTGCAGTC 1589

1433 TGAGCCCACTGCTGGTCCTGGTGGTAACCGCTCTGTCCACCCTATTATCT 1482
     ||||| |||||| ||| ||| || | ||||| || | |||| ||||||
1590 TGAGCTCACTGCCGGTGCTGATGCTCACCGCCCTTGCTGCCCTGTTATCT 1639

1483 ......TTAACAGAAACATCATAGCTGCATTAAAAAAATACAATATGGAC 1526
          || |||||||| || |||||||||| ||| ||| ||||
1640 GTATCGTTGGCAGAAACGTCGTAGCTGCATCCGGGAAA.ACAGTATG... 1685

1527 ATGTAAAAAGACAAA....AACCAAGTTATCTGTTTCCTGTTCTCTTGTA 1572
         |||||||||| |||||||| |||| | |||||| |||||||||
1686 .....AAAAGACAAAAGAGAACCAAGTATTCTG.TCCCTGTCCTCTTGTA 1729

1573 TAGCTGAAATTCCAG.TTTAGGAGCTCAGTTGAGAAACAGTTCCATTCAA 1621
     || ||||||| ||||| |||| ||||| ||||||||| ||||| || |||
1730 TATCTGAAAATCCAGTTTTAAAAGCTCCGTTGAGAAGCAGTTCACCCAA 1779

1622 CTGGAACATTTTTTTTTTTTCCTTTTAAGAAAGCTTCTTGTGATCCTTCG 1671
     |||||||   | ||| || ||||||||||| |||||| ||| |
1780 CTGGAAC....TCTTTCCTTGTTTTTAAGAAAG...CTTGTGGCCCTCAG 1822

1672 GGGCTTCTGT 1681
     ||||||||||
1823 GGGCTTCTGT 1832
```

FIG. 3A-3

```
  1 MFLATLYFALPLLDLLLSAEVSGGDRLDCVKASDQCLKEQSCSTKYRTLR  50
    |||||||||||||||:|||||||||||||||||||||||||||||||||||
  1 MFLATLYFALPLLDLLMSAEVSGGDRLDCVKASDQCLKEQSCSTKYRTLR  50

 51 QCVAGKETNFSLASGLEAKDECRSAMEALKQKSLYNCRCKRGMKKEKNCL  100
    |||||||||||||.||||||||||||||||||||||||||||||||||||
 51 QCVAGKETNFSLTSGLEAKDECRSAMEALKQKSLYNCRCKRGMKKEKNCL  100

101 RIYWSMYQSLQGNDLLEDSPYEPVNSRLSDIFRVVPFIS.....VEHIPK  145
    ||||||||||||||||||||||||||||||||.|||||     ||||.|
101 RIYWSMYQSLQGNDLLEDSPYEPVNSRLSDIFRAVPFISDVFQQVEHISK  150

146 GNNCLDAAKACNLDDICKKYRSAYITPCTTSVSNDVCNRRKCHKALRQFF  195
    |||||||||||||.||||||||||||||:||:||||||||||||||||||
151 GNNCLDAAKACNLDDTCKKYRSAYITPCTTSMSNEVCNRRKCHKALRQFF  200

196 DKVPAKHSYGMLFCSCRDIACTERRRQTIVPVCSYEEREKPNCLNLQDSC  245
    |||||||||||||||||||||||||||||||||||||||:||||.|||||
201 DKVPAKHSYGMLFCSCRDIACTERRRQTIVPVCSYEERERPNCLSLQDSC  250

246 KTNYICRSRLADFFTNCQPESRSVSSCLKENYADCLLAYSGLIGTVMTPN  295
    |||||||||||||||||||||||||.||||||||||||||||||||||||
251 KTNYICRSRLADFFTNCQPESRSVSNCLKENYADCLLAYSGLIGTVMTPN  300

296 YIDSSSLSVAPWCDCSNSGNDLEECLKFLNFFKDNTCLKNAIQAFGNGSD  345
    |:|||||||||||||||||||||:|||||||||||||||||||||||||||
301 YVDSSSLSVAPWCDCSNSGNDLEDCLKFLNFFKDNTCLKNAIQAFGNGSD  350

346 VTVWQPAFPVQTTTATTTTALRVKNKPLGPAGSENEIPTHVLPPCANLQA  395
    ||:||||.|||||||||||:|||||||||||||||||||||||||||||||
351 VTMWQPAPPVQTTTATTTTAFRVKNKPLGPAGSENEIPTHVLPPCANLQA  400

396 QKLKSNVSGNTHLCISNGNYEKEGL.GASSHITTKSMAAPPSCGLSPLLV  444
    |||||||||.||||:|:::::|:||  |||||||||||||||||:||.| |
401 QKLKSNVSGSTHLCLSDSDFGKDGLAGASSHITTKSMAAPPSCSLSSLPV  450

445 LVVTALSTLL..SLTETS  460
    |::|||..|| ||.|||
451 LMLTALAALLSVSLAETS  468
```

FIG. 3B

METANEPHRIC
MESENCHYME CELL

URETERIC BUD CELL

c-RET
LIGAND

c-RET
RECEPTOR

TYROSINE
KINASE
CYTOPLASMIC
DOMAIN

GROWTH

c-RET-IgG
FUSION
PROTEIN

c-RET-ALKALINE PHOSPHATASE
FUSION PROTEIN

FIG. 4A

cDNA LIBRARY

TRANSFECT INTO
MAMMALIAN
CELLS

DETECT CELLS THAT
EXPRESS c-RET
LIGAND WITH THE
FUSION PROTEINS

c-RET-IgG
FUSION PROTEIN

c-RET
LIGAND

c-RET LIGAND
cDNA

c-RET-ALKALINE PHOSPHATASE
FUSION PROTEIN

FIG. 4B

DAY 14
EMBRYONIC KIDNEY IN RNA

↓ REVERSE
   TRANSCRIPTASE
↓ TAG POLYMENASE
   WITH Kid-013 AND Kid-015
↓ DIGEST WITH Not1
   AND Sal1

Not1 ─────────────────── Sal1
RAT c-RET EXTRACELLDAM
├──────── 1919 bp ────────┤

Sal1 ─── Fc ─── Not1
693bp
pSAB144

DIGEST WITH Sal1
AND Not1
PURIFY 693bp FRAGMENT

Sal1 ─── Fc ─── Not1

DIGEST WITH
Not1

Not1 ─────── Sal1

PCR FRAGMENT
ENCODING TGF.,
RECEPTOR EXTRA-
CELLULAR DOMAIN

Not1 Sal1
CMV
PROMOTER
pSAB132

DIGEST
WITH Not1
AND Sal1

Sal1
Not1   Fc   Not1
CMV   2-4   Sal1

Not1 ─── RAT ─── Sal1
c-RET ED
pJCO11

DIGEST WITH Sal1
ISOLATE 700bp
Sal1 FRAGMENT CONTAINING
Fc DOMAIN

↓ DIGEST WITH Sal1

Sal1 ─── Fc ─── Sal1
         Not1

LIGATE

Sal1
Not1   RAT   Fc
c-RET ED
pJCO11   Not1
              Sal1
SAB132
BACKBONE
CMV

ISOLATE 2612bp
Not1 FRAGMENT
FROM pJC012.

CLONE INTO
Not1 SITE OF
pMDR901

Sal1
Not1   RAT   Fc
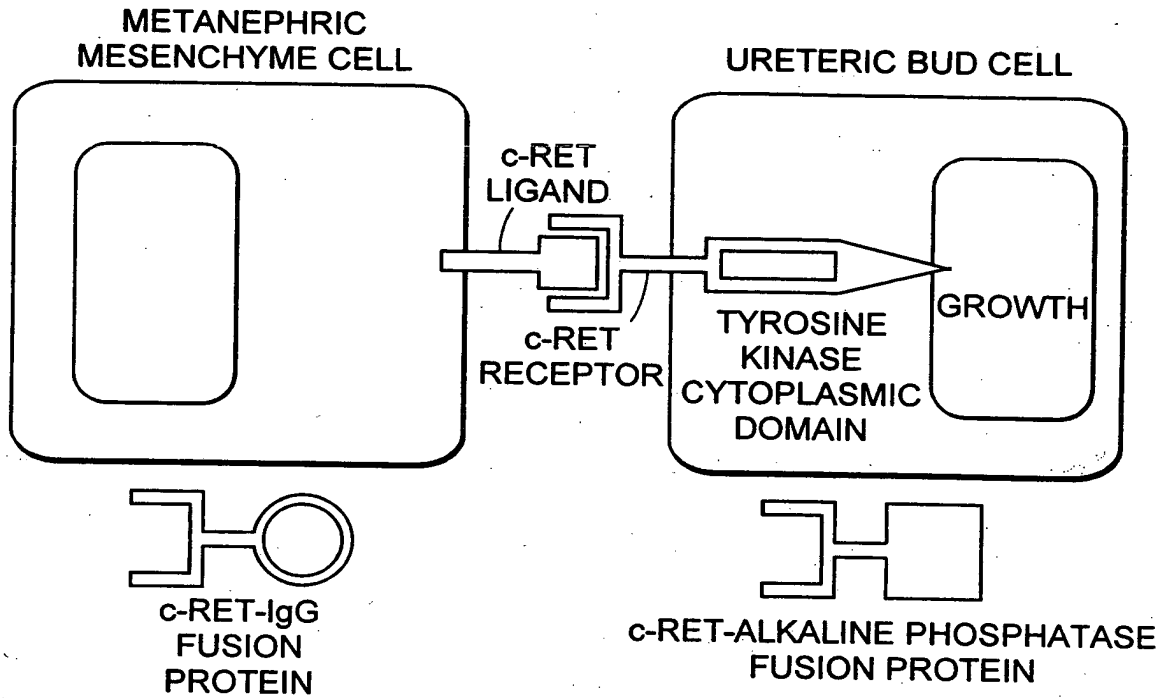c-RET ED
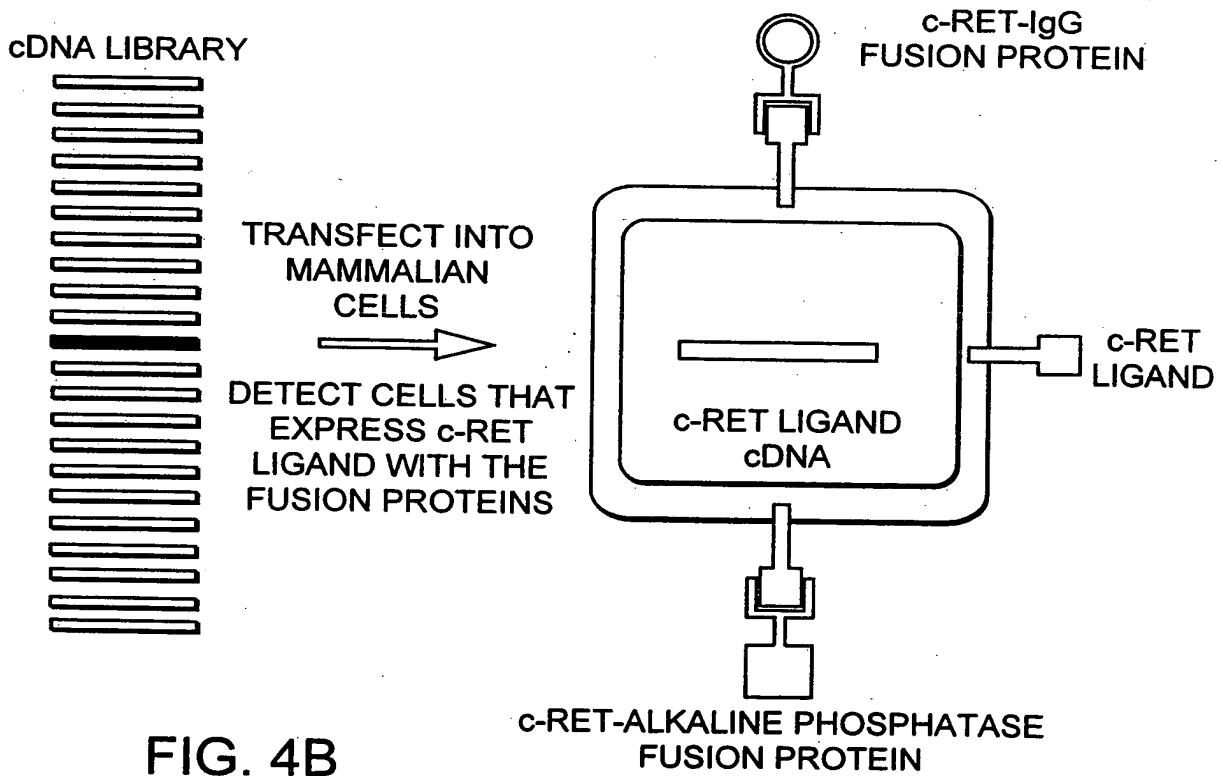pJCO22   Not1
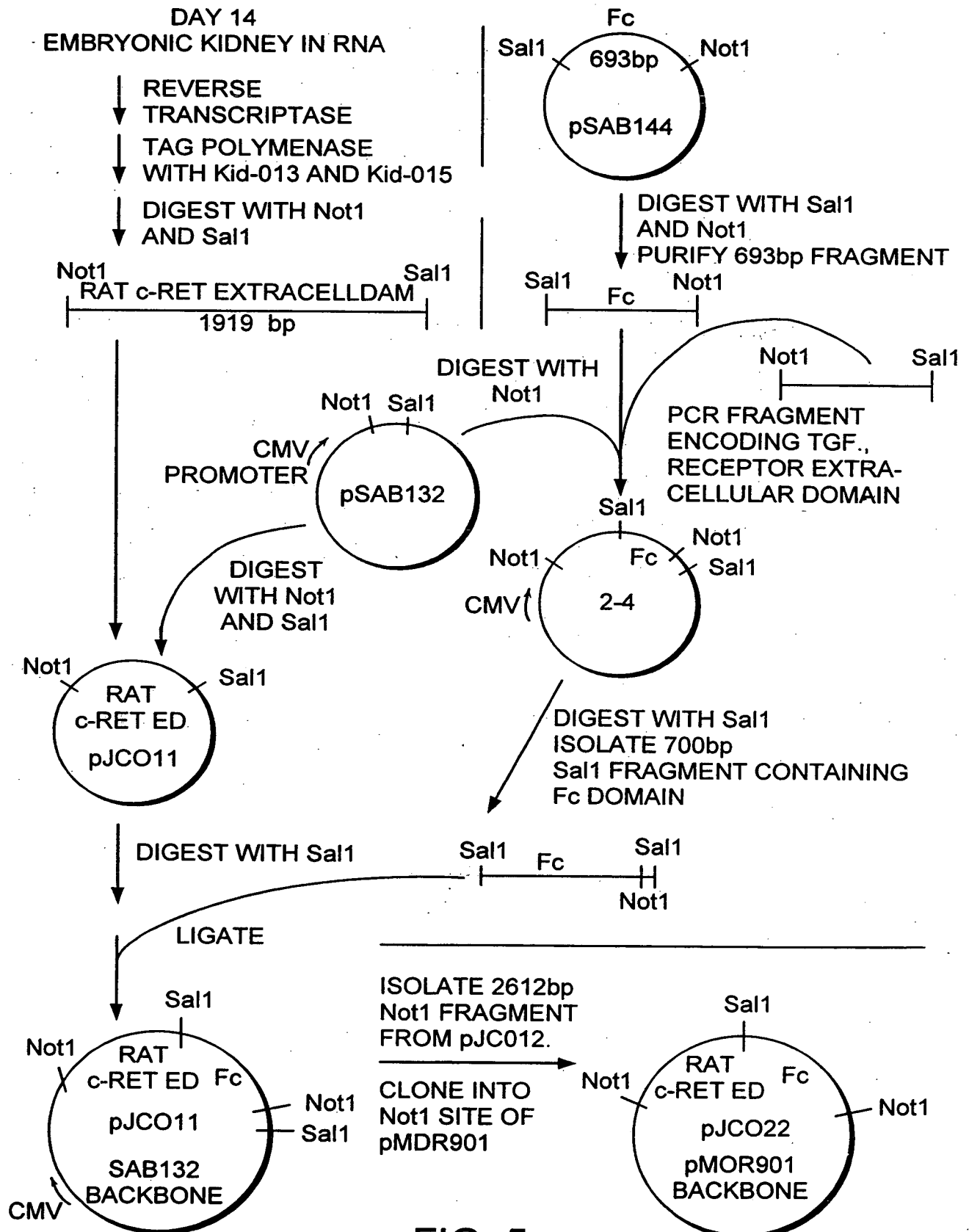pMOR901
BACKBONE

FIG. 5
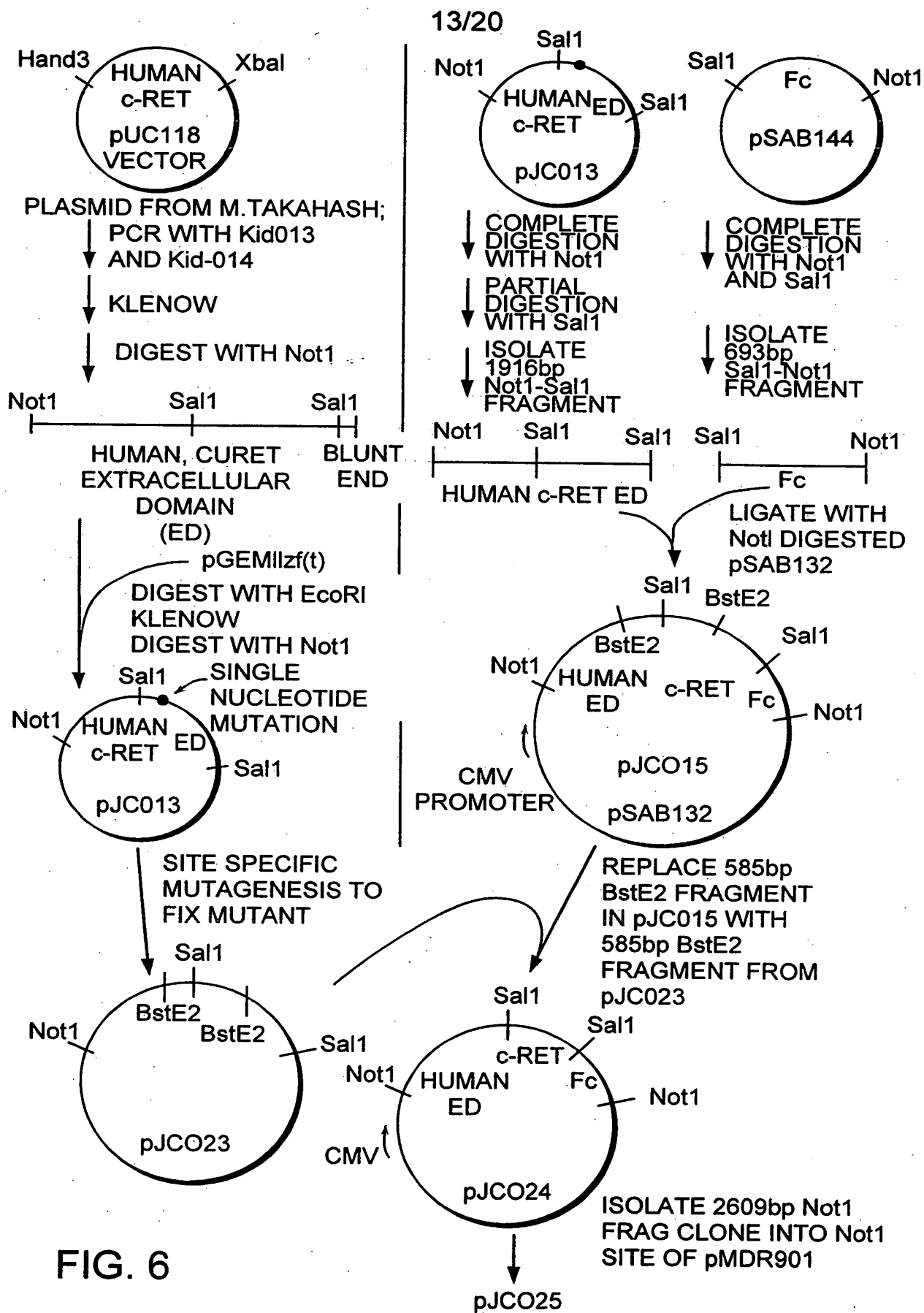
FIG. 6

```
  1  AAAAAACGGTGGGATTTATTTAACATGATCTTGGCAAACGTCTTCTGCCT   50
                      M  I  L  A  N  V  F  C  L

 51  CTTCTTCTTTCTAGACGAGACCCTCCGCTCTTTGGCCAGCCCTTCCTCCC  100
      F  F  F  L  D  E  T  L  R  S  L  A  S  P  S  S  L

101  TGCAGGGCCCCGAGCTCCACGGCTGGCGCCCCCCAGTGGACTGTGTCCGG  150
      Q  G  P  E  L  H  G  W  R  P  P  V  D  C  V  R

151  GCCAATGAGCTGTGTGCCGCCGAATCCAACTGCAGCTCTCGCTACCGCAC  200
      A  N  E  L  C  A  A  E  S  N  C  S  S  R  Y  R  T

201  TCTGCGGCAGTGCCTGGCAGGCCGCGACCGCAACACCATGCTGGCCAACA  250
      L  R  Q  C  L  A  G  R  D  R  N  T  M  L  A  N  K

251  AGGAGTGCCAGGCGGCCTTGGAGGTCTTGCAGGAGAGCCCGCTGTACGAC  300
      E  C  Q  A  A  L  E  V  L  Q  E  S  P  L  Y  D

301  TGCCGCTGCAAGCGGGGCATGAAGAAGGAGCTGCAGTGTCTGCAGATCTA  350
      C  R  C  K  R  G  M  K  K  E  L  Q  C  L  Q  I  Y

351  CTGGAGCATCCACCTGGGGCTGACCGAGGGTGAGGAGTTCTACGAAGCCT  400
      W  S  I  H  L  G  L  T  E  G  E  E  F  Y  E  A  S

401  CCCCCTATGAGCCGGTGACCTCCCGCCTCTCGGACATCTTCAGGCTTGCT  450
      P  Y  E  P  V  T  S  R  L  S  D  I  F  R  L  A

451  TCAATCTTCTCAGGGACAGGGGCAGACCCGGTGGTCAGCGCCAAGAGCAA  500
      S  I  F  S  G  T  G  A  D  P  V  V  S  A  K  S  N

501  CCATTGCCTGGATGCTGCCAAGGCCTGCAACCTGAATGACAACTGCAAGA  550
      H  C  L  D  A  A  K  A  C  N  L  N  D  N  C  K  K

551  AGCTGCGCTCCTCCTACATCTCCATCTGCAACCGCGAGATCTCGCCCACC  600
      L  R  S  S  Y  I  S  I  C  N  R  E  I  S  P  T

601  GAGCGCTGCAACCGCCGCAAGTGCCACAAGGCCCTGCGCCAGTTCTTCGA  650
      E  R  C  N  R  R  K  C  H  K  A  L  R  Q  F  F  D

651  CCGGGTGCCCAGCGAGTACACCTACCGCATGCTCTTCTGCTCCTGCCAAG  700
      R  V  P  S  E  Y  T  Y  R  M  L  F  C  S  C  Q  D

701  ACCAGGCGTGCGCTGAGCGCCGCCGGCAAACCATCCTGCCCAGCTGCTCC  750
      Q  A  C  A  E  R  R  R  Q  T  I  L  P  S  C  S

751  TATGAGGACAAGGAGAAGCCCAACTGCCTGGACCTGCGTGGCGTGTGCCG  800
      Y  E  D  K  E  K  P  N  C  L  D  L  R  G  V  C  R

801  GACTGACCACCTGTGTCGGTCCCGGCTGGCCGACTTCCATGCCAATTGTC  850
      T  D  H  L  C  R  S  R  L  A  D  F  H  A  N  C  R
```

FIG. 7a

```
 851  GAGCCTCCTACCAGACGGTCACCAGCTGCCCTGCGGACAATTACCAGGCG   900
        A   S   Y   Q   T   V   T   S   C   P   A   D   N   Y   Q   A

 901  TGTCTGGGCTCTTATGCTGGCATGATTGGGTTTGACATGACACCTAACTA   950
        C   L   G   S   Y   A   G   M   I   G   F   D   M   T   P   N   Y

 951  TGTGGACTCCAGCCCCACTGGCATCGTGGTGTCCCCCTGGTGCAGCTGTC  1000
        V   D   S   S   P   T   G   I   V   V   S   P   W   C   S   C   R

1001  GTGGCAGCGGGAACATGGAGGAGGAGTGTGAGAAGTTCCTCAGGGACTTC  1050
        G   S   G   N   M   E   E   E   C   E   K   F   L   R   D   F

1051  ACCGAGAACCCATGCCTCCGGAACGCCATCCAGGCCTTTGGCAACGGCAC  1100
        T   E   N   P   C   L   R   N   A   I   Q   A   F   G   N   G   T

1101  GGACGTGAACGTGTCCCCAAAAGGCCCCTCGTTCCAGGCCACCCAGGCCC  1150
        D   V   N   V   S   P   K   G   P   S   F   Q   A   T   Q   A   P

1151  CTCGGGTGGAGAAGACGCCTTCTTTGCCAGATGACCTCAGTGACAGTACC  1200
        R   V   E   K   T   P   S   L   P   D   D   L   S   D   S   T

1201  AGCTTGGGGACCAGTGTCATCACCACCTGCACGTCTGTCCAGGAGCAGGG  1250
        S   L   G   T   S   V   I   T   T   C   T   S   V   Q   E   Q   G

1251  GCTGAAGGCCAACAACTCCAAAGAGTTAAGCATGTGCTTCACAGAGCTCA  1300
        L   K   A   N   N   S   K   E   L   S   M   C   F   T   E   L   T

1301  CGACAAATATCATCCCAGGGAGTAACAAGGTGATCAAACCTAACTCAGGC  1350
        T   N   I   I   P   G   S   N   K   V   I   K   P   N   S   G

1351  CCCAGCAGAGCCAGACCGTCGGCTGCCTTGACCGTGCTGTCTGTCCTGAT  1400
        P   S   R   A   R   P   S   A   A   L   T   V   L   S   V   L   M

1401  GCTGAAACTGGCCTTGTAGGCTGTGGGAACCGAGTCAGAAGATTTTTGAA  1450
        L   K   L   A   L

1451  AGCTACGCAGACAAGAACAGCCGCCTGACGAAATGGAAACACACACAGAC  1500

1501  ACACACACACCTTGCAAAAAAAAATTGTTTTTCCCACCTTGTCGCTGAA  1550

1551  CCTGTCTCCTCCCAGGTTTCTTCTCTGGAGAAGTTTTTGTAAACCAAACA  1600

1601  GACAAGCAGGCAGGCAGCCTGAGAGCTGGCCCAGGGGTCCCCTGGCAGGG  1650

1651  GAAACTCTGGTGCCGGGGAGGGCACGAGGCTCTAGAAATGCCCTTCACTT  1700

1701  TCTCCTGGTGTTTTTCTCTGGACCCTTCTGAAGCAGAGACCGGACAAG  1750

1751  AGCCTGCAGCGGAAGGGACTCTGGGCTGTGCCTGAGGCTGGCTGGGGGCA  1800

1801  GGACAACACAGCTGCTTCCCCAGGCTGCCCACTCTGGGGACCCGCTGGGG  1850

1851  GCTGGCAGAGGGCATCGGTCAGCGGGGCAGCGGGGCTG  1888
```

FIG. 7b

```
  1 MILANVFCLFFFLDETLRSLASPSSLQGPELHGWRPPVDCVRANELCAAE  50
    |:||.::   : :||  |...|.:.            .:|||:|.: |  |
  1 MFLATLYFALPLLDLLLSAEVSGGD...........RLDCVKASDQCLKE  39

 51 SNCSSRYRTLRQCLAGRDRN.....TMLANKECQAALEVLQESPLYDCRC  95
    .||.:|||||||:||::  |     .: |..||..|:|.|.:..||:|||
 40 QSCSTKYRTLRQCVAGKETNFSLASGLEAKDECRSAMEALKQKSLYNCRC  89

 96 KRGMKKELQCLQIYWSIHLGLTEGEEFYEASPYEPVTSRLSDIFRLASIF 145
    ||||||| .||.||||:. :| :|:::.|.|||||||.||||||||||.
 90 KRGMKKEKNCLRIYWSMYQSL.QGNDLLEDSPYEPVNSRLSDIFRVVPFI 138

146 SGTGADPVVSAKSNHCLDAAKACNLNDNCKKLRSSYISICNREISPTERC 195
    |..       :|:|:|||||||||||||:| |||.||.||. | .:| .: |
139 SVEHI.....PKGNNCLDAAKACNLDDICKKYRSAYITPCTTSVS.NDVC 182

196 NRRKCHKALRQFFDRVPSEYTYRMLFCSCQDQACAERRRQTILPSCSYED 245
    ||||||||||||||:||....| ||||||.| ||.|||||||:| ||||:
183 NRRKCHKALRQFFDKVPAKHSYGMLFCSCRDIACTERRRQTIVPVCSYEE 232

246 KEKPNCLDLRGVCRTDHLCRSRLADFHANCRASYQTVTSCPADNYQACLG 295
    :||||||:|.: |:|:..:|||||||| .||.:. ..|.|| :||..||
233 REKPNCLNLQDSCKTNYICRSRLADFFTNCQPESRSVSSCLKENYADCLL 282

296 SYAGMIGFDMTPNYVDSSPTGIVVSPWCSCRGSGNMEEECEKFLRDFTEN 345
    .|.|:||  |||||:||| :: |.|||.|..||| ||| |||. |.:|
283 AYSGLIGTVMTPNYIDSS..SLSVAPWCDCSNSGNDLEECLKFLNFFKDN 330

346 PCLRNAIQAFGNGTDVNVSPKGPSFQATQAPRVEKTPSLPDDLSDSTS.. 393
    .||:||||||||||.||.|...: .|.| | |        ..|:...|
331 TCLKNAIQAFGNGSDVTVWQPAFPVQTTTATTTTALRVKNKPLGPAGSEN 380

394 .LGTSVITTCTSVQEGLKANNS..KELSMCFTELTTNIIPGSNKVIKPN 440
    :.|  |:..|..:|.| ||.| |  ..|:::  .:....: :.:|..:....
381 EIPTHVLPPCANLQAQKLKSNVSGNTHLCISNGNYEKEGLGASSHITTKS 430

441 SGPSRARPSAALTVLSVLMLKLAL 464
    ::... . .:| ||| | |. |
431 MAAPPSCGLSPLLVLVVTALSTLL 454
```

FIG. 8

```
  1    CGCAGGCAGAGCGCTGTCGCATCCCGGGCGTCCACCCGCCATGGGGCTCT     50
                                                 M  G  L  S

 51    CCTGGAGCCCGCGACCTCCACTGCTGATGATCCTGCTACTGGTGCTGTCG    100
        W  S  P  R  P  P  L  L  M  I  L  L  L  V  L  S

101    TTGTGGCTGCCACTTGGAGCAGGAAACTCCCTTGCCACAGAGAACAGGTT    150
        L  W  L  P  L  G  A  G  N  S  L  A  T  E  N  R  F

151    TGTGAACAGCTGTACCCAGGCCAGAAAGAAATGCGAGGCTAATCCCGCTT    200
        V  N  S  C  T  Q  A  R  K  K  C  E  A  N  P  A  C

201    GCAAGGCTGCCTACCAGCACCTGGGCTCCTGCACCTCCAGTTTAAGCAGG    250
        K  A  A  Y  Q  H  L  G  S  C  T  S  S  L  S  R

251    CCGCTGCCCTTAGAGGAGTCTGCCATGTCTGCAGACTGCCTAGAGGCAGC    300
        P  L  P  L  E  E  S  A  M  S  A  D  C  L  E  A  A

301    AGAACAACTCAGGAACAGCTCTCTGATAGACTGCAGGTGCCATCGGCGCA    350
        E  Q  L  R  N  S  S  L  I  D  C  R  C  H  R  R  M

351    TGAAGCACCAAGCTACCTGTCTGGACATTTATTGGACCGTTCACCCTGCC    400
        K  H  Q  A  T  C  L  D  I  Y  W  T  V  H  P  A

401    CGAAGCCTTGGTGACTACGAGTTGGATGTCTCACCCTATGAAGACACAGT    450
        R  S  L  G  D  Y  E  L  D  V  S  P  Y  E  D  T  V

451    GACCAGCAAACCCTGGAAAATGAATCTTAGCAAGTTGAACATGCTCAAAC    500
        T  S  K  P  W  K  M  N  L  S  K  L  N  M  L  K  P

501    CAGACTCGGACCTCTGCCTCAAATTTGCTATGCTGTGTACTCTTCACGAC    550
        D  S  D  L  C  L  K  F  A  M  L  C  T  L  H  D

551    AAGTGTGACCGCCTGCGCAAGGCCTACGGGGAGGCATGCTCAGGGATCCG    600
        K  C  D  R  L  R  K  A  Y  G  E  A  C  S  G  I  R

601    CTGCCAGCGCCACCTCTGCCTAGCCCAGCTGCGCTCCTTCTTTGAGAAGG    650
        C  Q  R  H  L  C  L  A  Q  L  R  S  F  F  E  K  A

651    CAGCAGAGTCCCACGCTCAGGGTCTGCTGCTGTGTCCCTGTGCACCAGAA    700
        A  E  S  H  A  Q  G  L  L  L  C  P  C  A  P  E

701    GATGCGGGCTGTGGGGAGCGGCGGCGTAACACCATCGCCCCCAGTTGCGC    750
        D  A  G  C  G  E  R  R  R  N  T  I  A  P  S  C  A

751    CCTGCCTTCTGTAACCCCCAATTGCCTGGATCTGCGGAGCTTCTGCCGTG    800
        L  P  S  V  T  P  N  C  L  D  L  R  S  F  C  R  A

801    CGGACCCTTTGTGCAGATCACGCCTGATGGACTTCCAGACCCACTGTCAT    850
        D  P  L  C  R  S  R  L  M  D  F  Q  T  H  C  H

851    CCTATGGACATCCTTGGGACTTGTGCAACTGAGCAGTCCAGATGTCTGCG    900
        P  M  D  I  L  G  T  C  A  T  E  Q  S  R  C  L  R
```

FIG. 9a

```
 901  GGCATACCTGGGGCTGATTGGGACTGCCATGACCCCAAACTTCATCAGCA   950
       A  Y  L  G  L  I  G  T  A  M  T  P  N  F  I  S  K

 951  AGGTCAACACTACTGTTGCCTTAAGCTGCACCTGCCGAGGCAGCGGCAAC  1000
        V  N  T  T  V  A  L  S  C  T  C  R  G  S  G  N

1001  CTACAGGACGAGTGTGAACAGCTGGAAAGGTCCTTCTCCCAGAACCCCTG  1050
       L  Q  D  E  C  E  Q  L  E  R  S  F  S  Q  N  P  C

1051  CCTCGTGGAGGCCATTGCAGCTAAGATGCGTTTCCACAGACAGCTCTTCT  1100
       L  V  E  A  I  A  A  K  M  R  F  H  R  Q  L  F  S

1101  CCCAGGACTGGGCAGACTCTACTTTTTCAGTGGTGCAGCAGCAGAACAGC  1150
        Q  D  W  A  D  S  T  F  S  V  V  Q  Q  Q  N  S

1151  AACCCTGCTCTGAGACTGCAGCCCAGGCTACCCATTCTTTCTTTCTCCAT  1200
       N  P  A  L  R  L  Q  P  R  L  P  I  L  S  F  S  I

1201  CCTTCCCTTGATTCTGCTGCAGACCCTCTGGTAGCTGGGCTTCCTCAGGG  1250
        L  P  L  I  L  L  Q  T  L  W

1251  TCCTTTGTCCTCTCCACCACACCCAGACTGATTTGCAGCCTGTGGTGGGA  1300

1301  GAGAACTCGCCAGCCTGTGGAAGAAGACGCAGCGTGCTACACAGCAACCC  1350

1351  GGAACCAACCAGGCATTCCGCAGCACATCCCGTCTGCTCCAGAAGAGGTC  1400

1401  TTAGAAGTGAGGGCTGTGACCCTTCCGATCCTGAGCGGCTAGTTTTCAAA  1450

1451  CCTCCCTTGCCCCTGCTTCCTTCTGGCTCAGGCTGCTCCTCCTTAGGACT  1500

1501  TTGTGGGTCCAGTTTTGCCTTCTGTTCTGATGGTGATTAGCGGCTCACCT  1550

1551  CCAGCGCTTCTTCCTGTTTCCCAGGACCACCCAGAGGCTAAGGAATCAGT  1600

1601  CATTCCCTGTTGCCTTCTCCAGGAAGGCAGGCTAAGGGTTCTGAGGTGAC  1650

1651  TGAGAAAAATGTTTCCTTTGTGTGGAAGGCTGGTGCTCCAGCCTCCACGT  1700

1701  CCCTCTGAATGGAAGATAAAAACCTGCTGGTGTCTTGACTGCTCTGCCAG  1750

1751  GCAATCCTGAACATTTGGGCATGAAGAGCTAAAGTCTTTGGGTCTTGTTT  1800

1801  AACTCCTATTACTGTCCCCAAATTCCCTAGTCCCTTGGGTCATGATTAA   1850

1851  ACATTTTGACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA  1889
```

FIG. 9b

```
  1   TGTGGACGCGCGCTTCGGAGTTGGAGGGCGGCGCCCAGGACCCTGGTGGG    50

 51   AGAGTGTGTGCGTCGCGCTGGAGGGCGGGAGGCGGGGGCGGGAGGTGCCG   100

101   GTCGAGGGAGCCCCGCTCTCAGAGCTCCAGGGGAGGAGCGAGGGGAGCGC   150

151   GGAGCCCGGCGCCTACAGCTCGCCATGGTGCGCCCCCTGAACCCGCGACC   200
                                  M  V  R  P  L  N  P  R  P

201   GCTGCCGCCCGTAGTCCTGATGTTGCTGCTGCTGCTGCCGCCGTCGCCGC   250
       L  P  P  V  V  L  M  L  L  L  L  P  P  S  P  L

251   TGCCTCTCGCAGCCGGAGACCCCCTTCCCACAGAAAGCCGACTCATGAAC   300
       P  L  A  A  G  D  P  L  P  T  E  S  R  L  M  N

301   AGCTGTCTCCAGGCCAGGAGGAAGTGCCAGGCTGATCCCACCTGCAGTGC   350
       S  C  L  Q  A  R  R  K  C  Q  A  D  P  T  C  S  A

351   TGCCTACCACCACCTGGATTCCTGCACCTCTAGCATAAGCACCCCACTGC   400
       A  Y  H  H  L  D  S  C  T  S  S  I  S  T  P  L  P

401   CCTCAGAGGAGCCTTCGGTCCCTGCTGACTGCCTGGAGGCAGCACAGCAA   450
       S  E  E  P  S  V  P  A  D  C  L  E  A  A  Q  Q

451   CTCAGGAACAGCTCTCTGATAGGCTGCATGTGCCACCGGCGCATGAAGAA   500
       L  R  N  S  S  L  I  G  C  M  C  H  R  R  M  K  N

501   CCAGGTTGCCTGCTTGGACATCTATTGGACCGTTCACCGTGCCCGCAGCC   550
       Q  V  A  C  L  D  I  Y  W  T  V  H  R  A  R  S  L

551   TTGGTAACTATGAGCTGGATGTCTCCCCCTATGAAGACACAGTGACCAGC   600
       G  N  Y  E  L  D  V  S  P  Y  E  D  T  V  T  S

601   AAACCCTGGAAAATGAATCTCAGCAAACTGAACATGCTCAAACCAGACTC   650
       K  P  W  K  M  N  L  S  K  L  N  M  L  K  P  D  S

651   AGACCTCTGCCTCAAGTTTGCCATGCTGTGTACTCTCAATGACAAGTGTG   700
       D  L  C  L  K  F  A  M  L  C  T  L  N  D  K  C  D

701   ACCGGCTGCGCAAGGCCTACGGGGAGGCGTGCTCCGGGCCCCACTGCCAG   750
       R  L  R  K  A  Y  G  E  A  C  S  G  P  H  C  Q

751   CGCCACGTCTGCCTCAGGCAGCTGCTCACTTTCTTCGAGAAGGCCGCCGA   800
       R  H  V  C  L  R  Q  L  L  T  F  F  E  K  A  A  E

801   GCCCCACGCGCAGGGCCTGCTACTGTGCCCATGTGCCCCCAACGACCGGG   850
       P  H  A  Q  G  L  L  L  C  P  C  A  P  N  D  R  G

851   GCTGCGGGGAGCGCCGGCGCAACACCATCGCCCCCAACTGCGCGCTGCCG   900
       C  G  E  R  R  R  N  T  I  A  P  N  C  A  L  P

901   CCTGTGGCCCCCAACTGCCTGGAGCTGCGGCGCCTCTGCTTCTCCGACCC   950
       P  V  A  P  N  C  L  E  L  R  R  L  C  F  S  D  P
```

FIG. 10a

```
 951  GCTTTGCAGATCACGCCTGGTGGATTTCCAGACCCACTGCCATCCCATGG  1000
       L  C  R  S  R  L  V  D  F  Q  T  H  C  H  P  M  D

1001  ACATCCTAGGAACTTGTGCAACAGAGCAGTCCAGATGTCTACGAGCATAC  1050
        I  L  G  T  C  A  T  E  Q  S  R  C  L  R  A  Y

1051  CTGGGGCTGATTGGGACTGCCATGACCCCCAACTTTGTCAGCAATGTCAA  1100
       L  G  L  I  G  T  A  M  T  P  N  F  V  S  N  V  N

1101  CACCAGTGTTGCCTTAAGCTGCACCTGCCGAGGCAGTGGCAACCTGCAGG  1150
        T  S  V  A  L  S  C  T  C  R  G  S  G  N  L  Q  E

1151  AGGAGTGTGAAATGCTGGAAGGGTTCTTCTCCACAACCCCTGCCTCACG  1200
        E  C  E  M  L  E  G  F  F  S  H  N  P  C  L  T

1201  GAGGCCATTGCAGCTAAGATGCGTTTTCACAGCCAACTCTTCTCCCAGGA  1250
        E  A  I  A  A  K  M  R  F  H  S  Q  L  F  S  Q  D

1251  CTGGCCACACCCTACCTTTGCTGTGATGGCACACCAGAATGAAAACCCTG  1300
        W  P  H  P  T  F  A  V  M  A  H  Q  N  E  N  P  A

1301  CTGTGAGGCCACAGCCCTGGGTGCCCTCTCTTTTCTCCTGCACGCTTCCC  1350
         V  R  P  Q  P  W  V  P  S  L  F  S  C  T  L  P

1351  TTGATTCTGCTCCTGAGCCTATGGTAGCTGGACTTCCCCAGGGCCCTCTT  1400
        L  I  L  L  L  S  L  W

1401  CCCCTCCACCACACCCAGGTGGACTTGCAGCCCACAAGGGGTGAGGAAAG  1450

1451  GACAGCAGCAGGAAGGAGGTGCAGTGCGCAGATGAGGGCACAGGAGAAGC  1500

1501  TAAGGGTTATGACCTCCAGATCCTTACTGGTCCAGTCCTCATTCCCTCCA  1550

1551  CCCCATCTCCACTTCTGATTCATGCTGCCCCTCCTTGGTGGCCACAATTT  1600

1601  AGCCATGTCATCTGGTGCCTGTGGGCCTTGCTTTATTCCTATTATTGTCC  1650

1651  TAAAGTCTCTCTGGGCTCTTGGATCATGATTAAACCTTTGACTTAAAAA  1699
```

FIG. 10b